Tensor-Product Kernel-based Representation encoding Joint MRI View Similarity

A. Álvarez-Meza*, D. Cárdenas-Peña*, A.E. Castro-Ospina*, M. Álvarez†, and G. Castellanos-Dominguez*

***Abstract*— To support 3D magnetic resonance image (MRI) analysis, a marginal image similarity (MIS) matrix holding MR inter-slice relationship along every axis view (Axial, Coronal, and Sagittal) can be estimated. However, mutual inference from MIS view information poses a difficult task since relationships between axes are nonlinear. To overcome this issue, we introduce a Tensor-Product Kernel-based Representation (TKR) that allows encoding brain structure patterns due to patient differences, gathering all MIS matrices into a single joint image similarity framework. The TKR training strategy is carried out into a low dimensional projected space to get less influence of voxel-derived noise. Obtained results for classifying the considered patient categories (gender and age) on real MRI database shows that the proposed TKR training approach outperforms the conventional voxel-wise sum of squared differences. The proposed approach may be useful to support MRI clustering and similarity inference tasks, which are required on template-based image segmentation and atlas construction.**

I. INTRODUCTION

Brain Magnetic Resonance (MR) imaging plays an important role in many medical applications like: *i*) identification of differences of functional brain structures along the time or space that can help to model evolution of pathologies, by instance, dementia, Alzheimer, and schizophrenia [1], *ii*) to develop realistic conductivity head models enhancing activity reconstruction accuracy [2], [3], *iii*) to extract spatial characteristics (as size, shape, and place) allowing to build representative anatomical models of populations [4].

All the applications enumerated above need accurate segmentation of brain regions and structures, which is fairly not easy to obtain due to image artifacts and low inter-structure contrast [5], [6]. To cope with these issues, template-based techniques have raised taking into account some prior spatial distributions of brain structure shapes. Specifically, priors are provided as a set of shape, intensity and/or functional models of structures as introduced in [7]. Nevertheless, segmentation quality is highly dependent on performed atlas-to-image registration. Besides, each obtained template may produce errors on accomplished segmentation results since ground-truth templates are manually drawn (not mentioning that a unimodal brain structure distribution shape is imposed) [8]. Therefore, to overcome those drawbacks, multi-atlas segmentation schemes have been proposed recently, where instead of just one, a set of atlases is registered to each query image. As a result, each voxel label is inferred over more information extracted from the atlas set so that influence of wrongly registered outliers or mislabeled templates is reduced. In this regard, head models are computed in [2] specifically for some subject-oriented demographic categories (as age, ethnicity, gender, and skull size), improving source localization by reducing the bias of anatomically unrepresentative atlases. The approach in [7] uses the presented mean shift algorithm perform atlas stratification to determine whether considered population is multi-modal and is best represented by an atlas per mode. Likewise, ranked atlas selection is performed in [8] by computing image similarities among subjects images based on measures like sums of squared differences, cross-correlation, and mutual information. Yet, convergence of both above introduced estimators can not be guaranteed in case of high-dimensional spaces.

Here, a kernel-based representation is proposed to improve MRI grouping analysis. To this end, inherent Inter-Slice Kernel (ISK) relationship is estimated to identify smooth MR slice variations that aim to highlight brain structure distributions. Besides, ISK-based feature representation is used to find each pairwise MR image marginal similarity (MIS), which are coupled by a Tensor-Product Kernel strategy to explore joint MRI similarity. The proposed approach is concretely developed to enhance both MRI data interpretability and separability using patient demographic information. Particularly, age and gender patient categories are studied. In this regard, our proposal is suitable to support MR image clustering and similarity measurement tasks required on template-based image segmentation and atlas construction.

II. TENSOR-PRODUCT KERNEL-BASED REPRESENTATION

Provided a MRI set $\mathcal{X}=\{\Psi_n \in \Omega\}$ to encode pair-wise image affinity, we propose a kernel-based representation as: $f(\Psi_n,\Psi_m)=\langle \varphi(\Psi_n), \varphi(\Psi_m)\rangle$, with $m,n=1,\ldots,N$; where $\varphi : \mathcal{X}\to\mathcal{H}$ is a nonlinear function mapping from the original feature space, $\Omega \subset \mathbb{R}^{L_a\times L_s\times L_c}$, to a Reproducing Kernel Hilbert Space, $\mathcal{H}\subset\mathbb{R}^q$. Though in practice it holds that $|\Omega| \ll q$ and $q\to\infty$, through the so called "*kernel trick*", there is no need for computing $\varphi(\cdot)$ [9].

We assume an original MR image space, Ω, holding a slice set $\Psi=\{\boldsymbol{X}_i^v : i=1,\ldots,L_v\}$ ordered according to the considered axis view, namely: Axial, Sagittal, and Coronal, respectively, noted as $v\in\{a,s,c\}$. Thus, Axial axis has L_a slices, each one represented by a matrix $\boldsymbol{X}_i^a \in \mathbb{R}^{L_s\times L_c}$.

*Universidad Nacional de Colombia, Sede Manizales, La Nubia, Colombia, Signal Processing and Recognition Group. {amalvarezme, dcardenasp, aecastroo, cgcastellanosd}@unal.edu.co

†Universidad Tecnológica de Pereira, Department of Electrical Engineering.malvarez@utp.edu.co

This work was supported by *Programa Nacional de Formación de Investigadores "Generación del Bicentenario"*, 2011/2012, *Programa Nacional de Jóvenes investigadores e innovadores*, 2012, and the research project 111056934461, all funded by COLCIENCIAS.

Further, we impose smooth variations between adjacent slices of Ψ to encode each pairwise Inter-Slice Kernel (ISK) relationship along the v-th axis into the matrix $\boldsymbol{S}^v = \{s_{ij}^v \in \mathbb{R}^+\}$, with $\boldsymbol{S}^v \in \mathbb{R}^{L_v \times L_v}$, where ISK matrix elements are calculated by a positive definite kernel function, $\mathcal{K}_1(\cdot,\cdot)$:

$$s_{ij}^v = \mathcal{K}_1\left(\boldsymbol{X}_i^v, \boldsymbol{X}_j^v\right). \tag{1}$$

Since ISK matrix allows representing high-dimensional image information along every axis, Marginal Image Similarity (MIS) can performed for each axis v based on pairwise comparison of image features (i,j) as follows:

$$k_{nm}^v = \mathcal{K}_2\left(\boldsymbol{S}_n^v, \boldsymbol{S}_m^v\right) \tag{2}$$

where $\mathcal{K}_2(\cdot,\cdot)$ is a positive definite and infinitely divisible kernel function producing the matrix $\boldsymbol{K}^v = \{k_{nm}^v \in \mathbb{R}^+\}$ with $\boldsymbol{K}^v \in \mathbb{R}^{N\times N}$.

Afterwards, to explore joint image similarity through all axes, we put forward a joint affinity measure of MIS matrices by introducing the following Tensor-Product Kernel Representation (TKR):

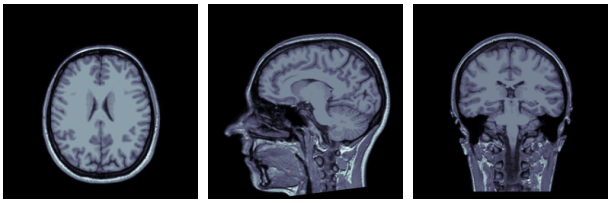
$$c_{nm} = \prod_{\forall v} \left(k_{nm}^v\right)^{\theta_v}, \tag{3}$$

where $\theta_v \in \mathbb{R}^+$ is the kernel exponent, $\boldsymbol{C} = \{c_{nm} \in \mathbb{R}^+\}$, and $\boldsymbol{C} \in \mathbb{R}^{N\times N}$. However, noisy MIS will lead to deleterious effects on the joint similarity in Eq. (3), due to $c_{nm} \to 0$ when $k_{nm}^v \to 0$. To cope with those effects, influence of the v-th MIS is decreased as $\theta_v \to 0$, so, $\left(k_{nm}^v\right)^{\theta_v} \to 1$. Besides, positive definite and infinitely divisible properties of kernels in Eq. (2) allow fixing arbitrary powers, θ_v, so that the resulting TKR in Eq. (3) is always positive definite.

III. EXPERIMENTS AND RESULTS

A. Demographic MRI Database

The IXI dataset is a brain imaging study holding MR images from 575 normal subjects aging between 20 and 80 years. Subjects were provided with T1, T2, PD, DTI, and angiogram volumes. All image sequences were obtained with three different scanners (Philips 1.5T, Philips 3T, and GE 3T), to be further anonymised and converted to NIFTI format. Additionally, basic demographic information for each subject is included (age, gender, ethnicity, among others). The whole dataset is publicly available online[1]. Since the current paper goal concerns atlas construction, only the T1 sequences were taken into account that were acquired with the GE 3T scanner for $N = 314$ subjects. Those T1 sequences are composed of $256 \times 256 \times 150$-sized volumes with a voxel size of $0.9375 \times 0.9375 \times 1.2\,mm$. Thus, the considered subset is composed of 139 male and 175 female subjects. Fig. 1 shows an example of the MR image for a concrete subject along three different views.



(a) Axial (b) Sagittal (c) Coronal

Fig. 1. Volume sample from the IXI database. Subject 002

B. Preprocessing over input MRI images

Two preprocessing steps are performed over the whole image dataset. Initially, each image is registered to the MNI305 template by an affine transform so that the whole dataset is referenced to the Talairach space [10]. Hence, each volume is re-sampled to $197 \times 233 \times 189$ size. Then, an intensity normalization procedure is performed by scaling each voxel value so that the mean intensity of the white matter is fixed to be 110 [11]. Both preprocessing steps, normalization and registering, are performed with the Freesurfer image analysis suite that is freely available online[2].

C. Learning Patient Patterns extracted from MRI Set

We explore two similarity-based image representation techniques to find patient patterns from MRI. The first one is a baseline where each image voxel is used as feature [8], while the second one uses the proposed TKR to encode each MRI pairwise relationship. Here, since all provided kernel functions are Gaussian, we introduce the following notation:

$$g\left(\mathrm{d}\left(z,z'\right);\sigma\right) \triangleq \exp\left(-\mathrm{d}\left(z,z'\right)^2/(2\sigma^2)\right),$$

where $\sigma \in \mathbb{R}^+$ is the kernel band-width; $z, z' \in \mathcal{Z}$ is a sample pair in a given feature space $\mathcal{Z}$, and $\mathrm{d}(\cdot,\cdot)$ is a distance operator in $\mathcal{Z}$.

As regards the voxel-wise approach, we calculate each element $y_{nm} \in \mathbb{R}^+$ of the MRI similarity matrix $\boldsymbol{Y} \in \mathbb{R}^{N\times N}$ using the Euclidean metric between n-th and m-th images as follows (Notation $\|\cdot\|_2$ stands for the 2-norm):

$$y_{nm} = g\left(||\operatorname{vec}(\Psi_n) - \operatorname{vec}(\Psi_m)||_2; \sigma_V\right), \tag{4}$$

notation $\operatorname{vec}(\cdot)$ stands for column-based concatenation of a given matrix. In turn, for the latter TKR approach, we calculate each ISK-based feature representation matrix $\boldsymbol{S}^v$ from MRI as in Eq. (1) using the Frobenious norm:

$$s_{ij}^v = g\left(||\boldsymbol{X}_i^v - \boldsymbol{X}_j^v||_F; \sigma_{s_v}\right). \tag{5}$$

As a result, we get three ISK matrices (one for each view) of size $L_v \in \{197, 233, 189\}$. Afterward, we calculate each MIS value encoding pairwise MRI relationship as in Eq. (5):

$$k_{nm}^v = g\left(||\boldsymbol{S}_n^v - \boldsymbol{S}_m^v||_F; \sigma_{k_v}\right). \tag{6}$$

It should be quoted that every kernel band-width in Eqs. (4)-(6) must be properly tuned. In the concrete case, we take into account that the variance of any Gaussian

[1] http://www.brain-development.org/

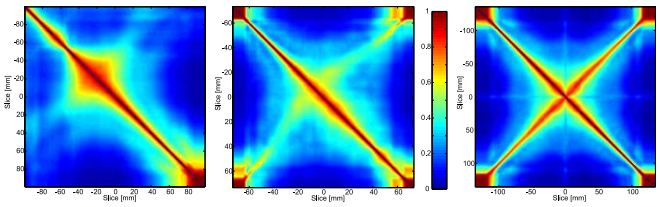
[2] http://surfer.nmr.mgh.harvard.edu/

kernel $g(\cdot;\sigma)$ tends to zero whenever σ tends to either zero or infinity. Therefore, to get an appropriate σ value spanning widely all similarity values, we propose to adjust the Gaussian kernel band-width employing the following criterion (Notation $\mathrm{var}(\cdot)$ stands for the variance operator):

$$\sigma^* = \arg\max_{\sigma} \left\{ \mathrm{var}(g(\cdot;\sigma) \right\},$$

In consequence, we obtain the following values in our experiments: $\sigma_V = 1280$, $\sigma_{s_1} = 133$, $\sigma_{s_2} = 129$, $\sigma_{s_3} = 130$, $\sigma_{k_1} = 9.25$, $\sigma_{k_2} = 7.02$, and $\sigma_{k_3} = 7.56$. Figure 2 shows an example of the proposed ISK representation. The red patches on the ISK representations are encoding all the MRI edges with no content, i.e., the background. As the Sagittal ISK (see Figure 2(c)) exhibits symmetry respect to the anti-diagonal, it is clear that the representation can read the head sagittal symmetry. Since the kernel shape varies as the brain structure distribution varies, it is inferred that proposed ISK encodes suitably the head shape. Although the latent phenomenon is the same for all ISK, each of them are providing a different view of the structure. Hence, aiming to include all the ISK information into a single kernel, proposed TKR is performed to estimate the MRI similarity matrix $\boldsymbol{C}$ coding axes view relations as presented in Eq. (3), where θ_v parameters are fixed by analyzing each kernel matrix $\boldsymbol{K}^v$ to accentuate MIS high variability that should be useful to identify MRI discriminative patterns as follows:

$$\theta_v = \frac{\mathrm{var}\left(\mathrm{vec}\left(\boldsymbol{K}^v\right)\right)}{\sum_v \mathrm{var}\left(\mathrm{vec}\left(\boldsymbol{K}^v\right)\right)}.$$



(a) Axial (b) Coronal (c) Sagittal

Fig. 2. ISK for a given image. All axes are measured in mm.

In our experiments, $\theta_1 = 0.32$, $\theta_2 = 0.35$, and $\theta_3 = 0.33$. Figure 3 presents the attained values of MRI similarity using TKR sorting them by gender and age *IXI* categories. As seen, the gender and age slots are highlighted from TKR supplying evidence about some possible MRI patterns.

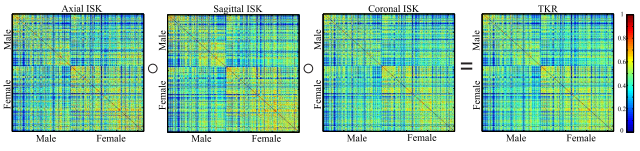


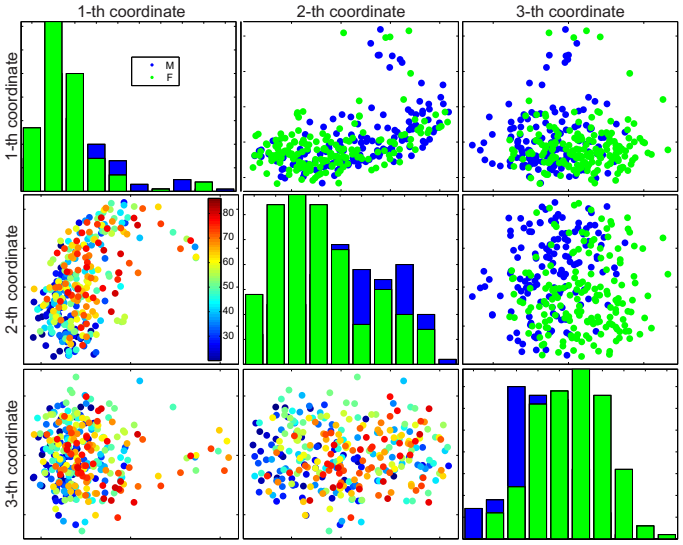
Fig. 3. MRI database TKR.

Furthermore, a low-dimensional space is computed from each matrix $\boldsymbol{Y}$ and $\boldsymbol{C}$ based on Kernel Principal Component Analysis (KPCA) to visually identify MRI clusters. Here, two *IXI* database categories are considered: gender and age. Therefore, the first three components for voxel-wise and TKR based representations are shown in Fig. 4. Regarding the age as a demographic category, by visual inspection of the first and second low-dimensional coordinates depicted in Fig. 4(b), it can be seen that the proposed methodology can unfold the age better than the baseline decomposition results (see Fig. 4(a)). Moreover, a quadratic dependence between second and first eigenvectors can be inferred. Additionally, a larger dispersion is shown on older subjects than on younger ones. This finding can be due to a larger head shape dispersion on older persons, which is according to anatomical head knowledge. It is known that brain anatomy is steady on middle age humans, while change (gray matter volume diminishes) faster on older humans [8]. With respect to gender category, the first and third low-dimensional coordinates seem to be suitable to distinguish *IXI* gender labels as seen in Figs. 4(a) and 4(b). Particularly, attained TKR-based projection allows to separate better between male and female labeled MRI than Voxel-wise one.

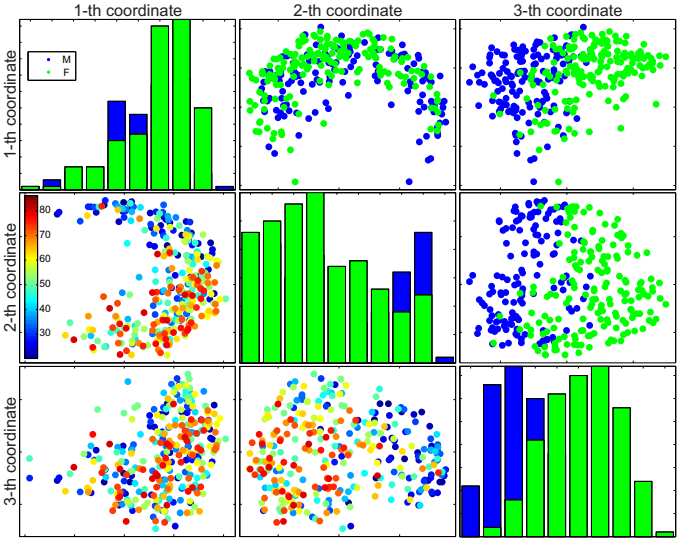
To verify the above mentioned statements, a k-nearest neighbor classifier is trained using each kernel-based induced distance from $\boldsymbol{Y}$ and $\boldsymbol{C}$. Table I presents the attained confusion matrices for a leave-one-out validation scheme, where four classes are considered: *adult male*, *senior male*, *adult female*, and *senior female*. Moreover, *adult* label corresponds to age values between 20 and 50 years, and *senior* label to age values higher than 50 years. Thus, our proposal attains higher classification accuracy than the baseline method. Indeed, presented ISK feature representation with TKR allows to identify similar brain structures by analyzing the joint axis view relationships. Regarding the Voxel-wise approach, it can obtain an acceptable gender discrimination, however, it is not able to distinguish age categories. So, complex brain structures, e.g., those related with patient age variations, can not be properly encode by the high-dimensional MR voxel representation.

IV. CONCLUSIONS

In this paper, a kernel-based image representation is introduced that is specifically devoted to support 3D MR image analysis. The proposed approach encodes smooth MRI inter-slice variations by a kernel-based function, which can be related to the brain structure distribution. Besides, ISK along each axis view, namely: Axial, Coronal, and Sagittal, are employed to estimate MRI marginal similarities as a tool to highlight head patterns. Furthermore, a Tensor-Product Kernel strategy is carried out over MIS to explore the joint MRI similarities for enhancing both data interpretability and separability using patient demographic information. Taking into account the attained results over a well-know MRI dataset, the proposed kernel-based representation methodology proved to find the natural inherent distributions of MR images, namely, age and gender categories. In addition, the proposed methodology improves data separability in comparison to state of the art algorithms based on Voxel-wise MR image representation. So, our proposal is suitable



(a) Voxel-wise projection



(b) TKR-based projection

Fig. 4. MRI database projections

to support MR image clustering and similarity measurement tasks required on template-based image segmentation.

As future work, three main research lines are proposed: *i*) Given that obtained decomposition vectors do not follow Gaussian distributions, other relaxed embedding methods, such as local linear embedding and Laplacian eigenmaps, will be tested aiming to improve the representation quality. *ii*) Supervised decomposition techniques will be proved to find representations suitable to distinguish other categories such as ethnicity or pathology subclasses. *iii*) The most straightforward research direction is to test our proposal as a template subset selector on MR image segmentation tasks so that the structure classification results are improved.

TABLE I

CONFUSION MATRICES [%]

Voxel-wise results

Category	*adult male*	*senior male*	*adult female*	*senior female*
adult male	91.66	55.22	06.94	04.85
senior male	0	11.94	0	01.94
adult female	06.94	02.98	83.33	27.18
senior female	01.39	29.85	09.72	66.01

Gender accuracy 88.85[%]
Gender-Age accuracy 65.61[%]

TKR results

Category	*adult male*	*senior male*	*adult female*	*senior female*
adult male	70.83	17.91	08.33	02.91
senior male	18.05	67.16	0	0
adult female	01.39	0	55.56	08.74
senior female	09.72	14.93	36.11	88.35

Gender accuracy 92.68[%]
Gender-Age accuracy 73.89[%]

REFERENCES

[1] Z. Lao, D. Shen, Z. Xue, B. Karacali, S. M. Resnick, and C. Davatzikos, "Morphological classification of brains via high-dimensional shape transformations and machine learning methods," *NeuroImage*, vol. 21, no. 1, pp. 46–57, Jan. 2004.

[2] P. Valdés-Hernández, N. von Ellenrieder, A. Ojeda-Gonzalez, S. Kochen, Y. Alemán-Gómez, C. Muravchik, and P. Valdés-Sosa, "Approximate average head models for EEG source imaging." *Journal of neuroscience methods*, vol. 185, no. 1, pp. 125–32, Dec. 2009.

[3] G. Strobbe, D. Cárdenas-Peña, V. Montes-Restrepo, P. Van Mierlo, G. Castellanos-Dominguez, and S. Vandenberghe, "Selecting volume conductor models for EEG source localization of epileptic spikes: preliminary results based on 4 operated epileptic patients," in *International Conference on Basic and Clinical multimodal Imaging*, 2013.

[4] A. Ericsson, P. Aljabar, and D. Rueckert, "Construction of a patient-specific atlas of the brain: Application to normal aging," in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Ieee, May 2008, pp. 480–483.

[5] U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI." *IEEE transactions on medical imaging*, vol. 26, no. 3, pp. 405–21, 2007.

[6] D. Cardenas-Pena, J. Martinez-Vargas, and G. Castellanos-Dominguez, "Local binary fitting energy solution by graph cuts for MRI segmentation." in *EMBC*, 2013.

[7] D. Blezek and J. Miller, "Atlas stratification." *Medical image analysis*, vol. 11, no. 5, pp. 443–57, 2007.

[8] P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy." *NeuroImage*, vol. 46, no. 3, pp. 726–38, 2009.

[9] B. Scholkopf and A. J. Smola, *Learning with Kernels*. Cambridge, MA, USA: The MIT Press, 2002.

[10] A. Evans, D. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, 1993, pp. 1813–1817.

[11] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale, "Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain." *Neuron*, vol. 33, no. 3, pp. 341–55, 2002.